Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

The cofactor-induced pre-active conformation in PhoB

Maria Solà,[a,b]* Devin L. Drew,[a]‡ Alexandre G. Blanco,[a,b] F. Xavier Gomis-Rüth[a] and Miquel Coll[a,b]

[a]Institut de Biologia Molecular de Barcelona, CSIC, Parc Científic de Barcelona, c/ Josep Samitier 1-5, 08028 Barcelona, Spain, and [b]Institut de Recerca Biomèdica, CSIC, Parc Científic de Barcelona, c/ Josep Samitier 1-5, 08028 Barcelona, Spain

‡ Present address: Department of Chemistry and Biochemistry, 339 NSB, University of California, San Diego, 9500 Gilman Drive, La Jolla, CA 92093-0375, USA.

Correspondence e-mail: msvcri@ibmb.csic.es

Received 29 March 2006
Accepted 26 June 2006

PDB Reference: PhoB, 2yin, r2yinsf.

PhoB is an *Escherichia coli* transcription factor from a two-component signal transduction system that is sensitive to limiting environmental phosphate conditions. It consists of an N-terminal receiver domain (RD) and a C-terminal DNA-binding domain. The protein is activated upon phosphorylation at the RD, an event that depends on Mg^{2+} binding. The structure of PhoB RD in complex with Mg^{2+} is presented, which shows three protomers in the asymmetric unit that interact across two different surfaces. One association is symmetric and has been described as a non-active dimerization contact; the other involves the $\alpha4$-$\beta5$-$\alpha5$ interface and recalls the contact found in activated PhoB. However, here this last interaction is not perfectly symmetric and helix $\alpha4$, which in the activated molecule undergoes a helical shift, becomes strongly destabilized in one of the interacting monomers. All protomers bind the cation in a similar manner but, interestingly, at the prospective binding site for the phosphate moiety the side chains of either Glu88 (in helix $\alpha4$) or Trp54 alternate and interact with active-site atoms. When Glu88 is inside the cavity, helix $\alpha4$ is arranged similarly to the unliganded wild-type structure. However, if Trp54 is present, the helix loses its contacts with the active-site cavity and vanishes. Accordingly, the presence of Trp54 in the active site induces a flexible state in helix $\alpha4$, potentially allowing a helical shift that phosphorylation would eventually stabilize.

1. Introduction

A large number of bacterial pathways depend on two-component signal transduction systems (2C-STS) to transmit extracellular or intracellular signals that activate cell metabolism. 2C-STSs govern chemotaxis, quorum sensing, nutrient uptake, nitrogen fixation, production of antibiotics, osmoregulation, sporulation, hormone-dependent developmental processes and pathogen invasion (Hoch, 2000; Stock *et al.*, 2000; Wang *et al.*, 2002; Wolanin *et al.*, 2002). 2C-STS are widespread in archaea, eubacteria, eukaryotic cell organelles and in yeast and plants, where these pathways are part of more complex signal transduction networks (Pao & Saier, 1997; Stock *et al.*, 2000; Forsberg *et al.*, 2001; Santos & Shiozaki, 2001; Grefen & Harter, 2004). Eubacteria containing 2C-STSs include pathogens, which are particularly able to adapt to changing environments. As such systems are absent in mammals, they can be envisaged as targets for the development of new antimicrobial drugs that will increase the medical armamentarium available in times of increasing bacterial antibiotic resistance (Barrett & Hoch, 1998; Stephenson & Hoch, 2002).

A typical 2C-STS system consists of a sensor component and a response regulator (RR) and may include additional regulatory components (Wick & Egli, 2004). The sensor is usually a membrane-spanning receptor consisting of an external ligand-binding domain, the sensor module and a cytosolic transmitter domain with histidine kinase activity. Indirect or direct binding of a ligand to the sensor triggers the transmission of the signal across the receptor molecule and hence across the membrane, resulting in phosphorylation of a specific conserved histidine in the cytosolic domain. The cognate RR subsequently directly or indirectly promotes a phosphotransfer from the sensor to a strictly conserved aspartate of its own receiver domain (RD). In most cases, the RR has additional domains, including a C-terminal effector domain (ED) whose activity is modulated by phosphorylation of the N-terminal RD. Phosphorylated RRs have a variety of responses, depending on the 2C-STS pathway. For example, the ED may bind to a cognate DNA sequence or the protein may change its oligomeric state and/or activate an enzymatic function. Dephosphorylation of the RR, either by a specific phosphatase or by the RR itself, switches off the system. Both phosphorylation and dephosphorylation are magnesium-dependent reactions (Lukat *et al.*, 1990; McCleary & Stock, 1994; Aizawa *et al.*, 2000; Stock *et al.*, 2000; Falke & Hazelbauer, 2001).

In *Escherichia coli*, the 229-amino-acid transcription factor PhoB RR comprises an N-terminal RD (residues 1–124) and a C-terminal ED with DNA-binding properties (residues 131–229) separated by a flexible six-residue linker. Together with its cognate sensor histidine kinase PhoR (Makino *et al.*, 1989), it is part of a 2C-STS that is sensitive to extracellular inorganic phosphate (P_i) levels (Wanner, 1996). In particular, PhoR detects changes in environmental P_i concentration either directly through its extracellular domain or through the high-affinity phosphate transporter (encoded by the *PstSCAB* genes), along with the ancillary protein PhoU. PhoR remains unphosphorylated while the organism grows in the presence of excess P_i. As soon as extracellular phosphate falls to starvation levels, PhoR autophosphorylates a conserved histidine residue from its histidine phosphorylation domain. The phosphate is subsequently transferred in a magnesium-dependent manner to a conserved aspartate residue (Asp53) within the N-terminal RD of PhoB. Upon phosphorylation, the PhoB ED binds with higher affinity to specific pho-box DNA sequences in the promoters of the PHO regulon, which comprises several genes and operons (McCleary, 1996). PhoB belongs to the OmpR/PhoB subfamily of RR, which are σ^{70} RNA polymerase subunit-dependent transcription factors. Transcription activation of the PHO regulon allows *E. coli* to bind environmental P_i with higher affinity and assimilate phosphorus from alternative sources such as pyrophosphate or metaphosphate (Wanner, 1993). To date, the three-dimensional structure of full-length PhoB has not been determined and therefore the structural details of the interaction between the domains are not known. Biochemical data demonstrate that the N-terminal RD inhibits the C-terminal ED DNA-binding activity. The removal of the RD results in a protein that binds DNA more tightly and activates transcription (Ellison & McCleary, 2000). Moreover, cross-linking studies on the full-length protein provided evidence that non-phosphorylated PhoB is able to form dimers and it was proposed that in the non-activated state there is a rapid equilibrium between the monomeric and dimeric forms that is shifted towards dimers by phosphorylation (McCleary, 1996). In contrast, functional studies of PhoB RD fused to the DNA-binding domain of the lambda repressor demonstrated that the RD only forms active dimers upon phosphorylation, suggesting that the PhoB ED prevents formation of the active dimer (Fiedler & Weiss, 1995). All this information led to two apparently contradictory scenarios: on one hand dimerization already happens before activation (McCleary, 1996), while on the other the inactive protein is a monomer owing to ED inhibition of the RD and only upon phosphorylation is the inhibition released and the protein becomes a dimer through the RD (Fiedler & Weiss, 1995). In this process, phosphorelay is meant to trigger interdomain rearrangements that release the inhibition and allow a dimer-mediated response.

We have previously determined the three-dimensional structures of the metal-free N-terminal domain of PhoB (referred to in the following as apo-PhoB^{R}), as well as that of the C-terminal domain either unbound or in complex with DNA. Apo-PhoB^{R} reveals strong structural similarity to RDs of the response regulator superfamily, but with a symmetric homodimeric quaternary arrangement not found in any orthologue (Solà *et al.*, 1998, 1999). The PhoB ED structure consists of a modified winged helix–turn–helix motif where DNA recognition occurs in tandem with an α-helix bound to the major groove and a wing interacting with the minor groove (Blanco *et al.*, 2002). The twofold apo-PhoB^{R} dimer positions the C-terminal ends of each RD at opposite sides, so that the two subsequent ED cannot simultaneously interact in tandem with a DNA molecule. In consequence, the apo-PhoB^{R} arrangement may correspond to the inactivated state of PhoB. In the present work, we report the structure of PhoB RD in complex with Mg^{2+} (PhoB^{RM}). This structure shows two different interfaces between the three protomers in the asymmetric unit, each involving a unique interaction surface. One is similar to the dimer interface in the apo-PhoB^{R} structure and the other is a small quasi-symmetric protein–protein interaction. The latter involves highly conserved residues and recalls a larger and symmetric interface found in the structure of PhoB RD activated (referred to in the following as PhoB^{RAc}) with the phosphoryl analogue beryllium fluoride (BeF_3^-; Bachhawat *et al.*, 2005). Interestingly, this second arrangement putatively places the two EDs close to each other at one edge of the dimer. Moreover, the inclusion of the cation in the active site results in the formation of a similar hydrogen-bonding network in all PhoB^{RM} protomers. However, the three active sites show heterogeneity at the specific area that in PhoB^{RAc} is occupied by BeF_3^- as, depending on the molecule in the crystal, it is alternatively filled by Trp54 or Glu88 side chains. In both cases the side chains interact with the active-site network, but the presence of Trp54 in the cavity correlates with a highly flexible

Table 1
Data-collection and refinement statistics for WT PhoB RD–Mg^{2+} complex.

Values in parentheses are for the last shell.

Wavelength (Å)	0.837
Space group	$C222_1$
Unit-cell parameters (Å)	$a = 50.7$, $b = 105.8$, $c = 135.5$
Estimated mosaicity (°)	0.75
No. of measurements	110494
No. of unique reflections	20372
Average multiplicity	5.4 (4.8)
Resolution range (Å)	45.2–2.08 (2.19–2.08)
$R_{\rm merge}$† (%)	7.6 (38.8)
Completeness (%)	91.4 (78.9)
$\langle I\rangle/\sigma(\langle I\rangle)$	5.2 (1.9)
Resolution range used for refinement (Å)	45.2–2.08
No. of reflections (working set/test set)	19679/658
R factor‡ (%)	23.1
Free R factor§ (%)	30.5
R.m.s.d. bonds (Å)	0.016
R.m.s.d. angles (°)	1.50
R.m.s.d. bonded B factors (Å^2)	1.58
No. of non-H protein atoms	2746
Water molecules	108
Metal ions	3 (Mg^{2+})
Average B factors (Å^2)	
Protein atoms	17.2
Solvent molecules	20.4
Metal ions	17.1

† $R_{\rm merge} = [\sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle|/\sum_{hkl}\sum_i I_i(hkl)] \times 100$, where $I_i(hkl)$ is the ith intensity measurement of reflection hkl and $\langle I(hkl)\rangle$ is its average intensity. ‡ R factor $= (\sum_{hkl} \big||F_{\rm obs}| - k|F_{\rm calc}|\big|/\sum_{hkl}|F_{\rm obs}|) \times 100$. § The free R factor is the R factor for a test set of reflections (>500) not used during the refinement.

conformation of helix α4, a segment at the surface involved in the quasi-symmetric interaction. Our results suggest that upon Mg^{2+} binding, this surface increases its flexibility, allowing the structural rearrangement of helix α4 needed for the active oligomerization.

2. Materials and methods

2.1. Protein expression and purification

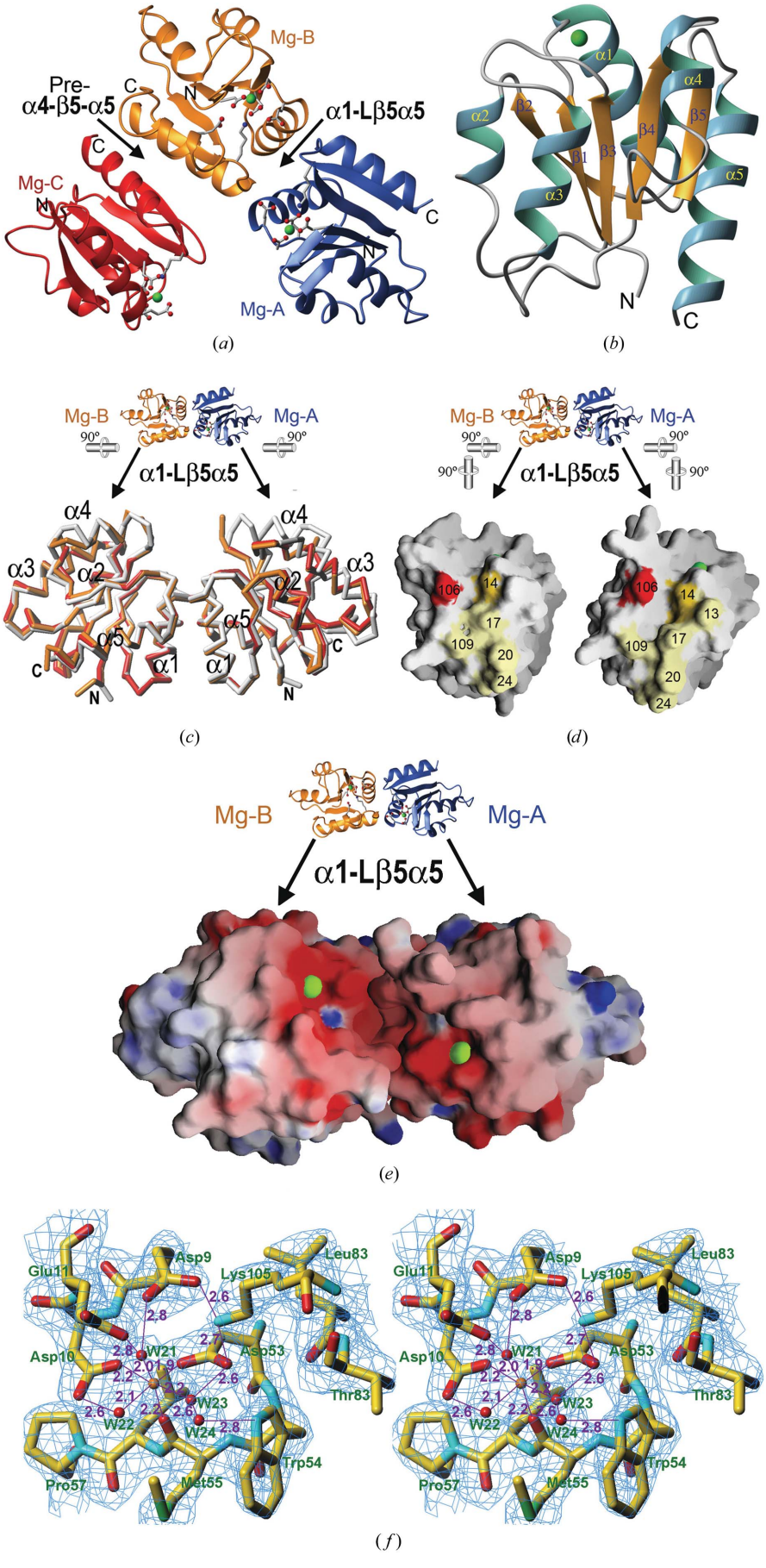
Native apo-PhoB RD was prepared and purified essentially as described elsewhere (Solà *et al.*, 1998, 1999). Briefly, a DNA fragment comprising the nucleic acid sequence encoding residues Met1–Ala127 of PhoB RD was subcloned into the pBAT-4 expression vector (Peränen *et al.*, 1996) and over-expressed in *E. coli* strain BL21 (DE3) at 298 K. The lysis buffer contained 20 m*M* Tris–HCl pH 8, which was used in subsequent steps of protein purification. The soluble fraction was purified in two consecutive steps of FPLC size-exclusion chromatography on a HiLoad 26/60 Superdex 75 column (Amersham Pharmacia). The sample was finally concentrated to 6 mg ml^{-1} and stored at 193 K.

2.2. Complex formation, protein crystallization, data collection and processing

Mg^{2+}-bound PhoB RD crystals were obtained by the hanging-drop vapour-diffusion method in Linbro crystallization plates from drops comprising 3 µl protein solution in the same buffer as used during studies with the unbound protein (Solà *et al.*, 1998) and 6 µl reservoir solution containing 20%(*w*/*v*) polyethylene glycol (PEG) 4000 (Fluka), 0.1 *M* Tris–HCl (Fluka) pH 8.0, 0.4 *M* magnesium acetate (J. T. Baker) and 0.02% sodium azide (J. T. Baker). These crystals belong to a new crystal form, space group $C222_1$, with unit-cell parameters $a = 50.7$, $b = 105.8$, $c = 135.5$ Å and three molecules per asymmetric unit ($V_{\rm M} = 2.1$ Å^3 Da^{-1}; 41% solvent content; Matthews, 1968). For comparison, apo-PhoB RD crystals belong to space group $P2_12_12_1$, with unit-cell parameters $a = 34.1$, $b = 60.4$, $c = 120.0$ Å and a dimer (WT-A and WT-B) in the asymmetric unit. In order to preserve crystals for transport and during data collection, they were flash-cryo-cooled in liquid nitrogen after soaking them in a solution containing 35% PEG 4000 for about 20 min. X-ray diffraction data were collected at 100 K on CCD area detectors at the ESRF synchrotron (Grenoble, France) beamlines ID14-EH2 and BM-16 and at beamline BW7B at DESY (Hamburg, Germany). Data were indexed and integrated with *MOSFLM* (Leslie, 1991) and scaled, merged and reduced with *SCALA* (Evans, 1993*a*). Data-collection and processing statistics are given in Table 1.

2.3. Structure solution and refinement

The structure of Mg^{2+}-bound PhoB RD was solved by Patterson search (Huber, 1965) with the *AMoRe* package (Navaza, 1994) using the apo-PhoB RD X-ray structure as a search model (PDB code 1b00; Solà *et al.*, 1999). Rotation and translation functions were calculated using data in the 15–4.0 Å resolution range. Firstly, a dimer as found in the apo structure was employed as a model and when its rotation and position had been elicited a second search using a monomer was undertaken to find the third molecule present in the crystallographic asymmetric unit. These calculations rendered two clear solutions, which showed a cumulative correlation coefficient in structure-factor amplitudes (defined in Navaza, 1994) of 55.2% (second highest unrelated 43.0%) and a crystallographic R factor of 43.5% (second highest unrelated 48.0%). The appropriately rotated and translated coordinates were refined using the maximum-likelihood criteria with *CNS* v.1.0 (Brünger *et al.*, 1998) and *REFMAC*5 (Murshudov *et al.*, 1997) employing bulk-solvent and anisotropic temperature-factor correction and using all data except for a sufficient number of reflections used for the cross-validation free R factor, which was monitored throughout during refinement. Refinement further included NCS restraints in the initial stages and TLS refinement as implemented in *REFMAC*5. The correctness of the solutions was confirmed by inspection of $\sigma_{\rm A}$-weighted $|2mF_{\rm obs} - dF_{\rm calc}|$ and $|mF_{\rm obs} - dF_{\rm calc}|$ Fourier maps, which showed clear density for deviating regions of the polypeptide chain and side chains, in particular around the metal-binding sites. After initial simulated-annealing refinement, cycles of positional refinement and temperature-factor refinement were alternated with manual model building using *TURBO-FRODO* (Carranza *et al.*, 1999) on a Silicon Graphics workstation. In the final steps of the refinement process, solvent molecules were placed manually at appropriate posi-



tions if a clear positive difference density was present at 2.5σ above average. The magnesium ions were identified on the basis of coordination geometry and ligand distances. The final Mg^{2+}-bound RD structure contains residues 3A–121A (excepting residues 87A–95A of helix α4), 1B–124B, 3C–122C (excepting residues 85C–96C of helix α4), three magnesium cations and 108 solvent molecules (see Table 1). Accordingly, a total of 39 amino acids (10%) are flexible and not defined by the electron density. This results in a high free R factor. A Ramachandran plot generated with *PROCHECK* (Laskowski *et al.*, 1993) shows all the residues in most favoured or in additional allowed regions except for

Figure 1
PhoB RD in complex with Mg^{2+}. (*a*) Richardson diagram of the three molecules present in the crystal asymmetric unit: Mg-A (blue), Mg-B (orange) and Mg-C (red). The magnesium ions are represented as green spheres and the coordinating protein residue side chains are shown as ball-and-stick models. (*b*) Richardson diagram of the PhoB RD molecule Mg-B. Helices are shown as blue ribbons and labelled (α1–α5) and β-strands as orange arrows (β1–β5). The positions of the N- and the C-termini are also depicted. The green sphere signifies the bound magnesium ion. (*c*) Superimposition of the C^{α} traces of the dimers linked through an α1-Lβ5α5 interface as found in WT-A/B (white), Mg-A/B (orange) and Mg-C/C′ (red). Note that helix α3 undergoes a rigid-body displacement on going from the apo protomer (on the right) to one of the magnesium-bound molecules. (*d*) The α1-Lβ5α5 interface made up by Mg-A (right, blue) and Mg-B (left, orange). The monomers are superimposed on their Connolly solid surface and have been rotated to show the surfaces in contact. The Mg^{2+} ions coordinated in the active site are depicted as green spheres. The residues engaged in homodimeric interactions are shown by their number and their percentage of conservation (see §2), ranging from 0% (pale yellow) to 100% (red). (*e*) Mg-A/B dimer superimposed with its Connolly solid surface coloured according to its electrostatic potential, from red ($-20k_BT/e$) to blue ($+20k_BT/e$). The view is the same as in Fig. 1(*a*), as indicated in the thumbnail. Note the continuous electronegative depression connecting both active sites, which contain Mg^{2+} ions shown as green spheres. (*f*) Stereo plot showing the final ($2mF_{obs} - dF_{calc}$) electron density (contoured at 1σ) superimposed with the final refined model around the cation-binding site of Mg-C. Intervening residues are displayed as sticks and labelled, as are participating solvent molecules (red spheres) and the magnesium cation (orange sphere). Hydrogen bonds and salt bridges are shown as violet lines and provided with a label indicating the distance (in Å) between the interacting atoms.

Glu96, which is at the end of helix $\alpha4$ in a poorly defined region. All three molecules show two *cis*-peptide bonds between residues 44 and 45 and 105 and 106, respectively. Final refinement statistics are shown in Table 1. The coordinates for the refined PhoB model in complex with magnesium have been deposited with the Protein Data Bank at http//www.ebi.ac.uk/msd, with accession code 2yin. During the deposition procedure some atoms/residues where renamed/renumbered by the PDB (see REMARKS section of the PDB entry).

2.4. Miscellaneous

Least-squares superimpositions were carried out with *LSQMAN* (Kleywegt & Jones, 1994). In order to calculate the distances between similar atoms of different monomers, only the cores (β-sheet strands) of the molecules were superimposed onto apo-PhoB RD molecule A (WT-A) and then compared. Dimer superimpositions were calculated using all C^{α} atoms. The conservation values of residues within the OmpR/PhoB family of RD were computed as follows. Firstly, the 30 most similar sequences to PhoB were identified with *PSI-BLAST* (Altschul & Koonin, 1998) and subjected to multiple alignment using *ClustalW* (both programs can be found at http://www.expasy.ch). The resulting sequence superposition was submitted to *ESPript* (Gouet, Courcelle *et al.*, 1999), which converted the similarity values to pseudo-temperature factors integrated into the PDB coordinates. Figures were produced with *GRASP* (Nicholls *et al.*, 1991), *SETOR* (Evans, 1993*b*), *MOLMOL* (Koradi *et al.*, 1996) and *TURBO-FRODO* (Carranza *et al.*, 1999). The interface area between molecules was calculated using *CNS*. Identification of the residues involved in the protomer interactions was performed by visual inspection of the interfaces. The corresponding stereochemical analysis was performed by submitting the coordinates to the protein–protein interactions server http://www.biochem.ucl.ac.uk/bsm/PP/server/index.html (Jones & Thornton, 1996).

3. Results

3.1. The interface of apo-PhoB RD (the '$\alpha1$-L$\beta5\alpha5$ interface') is not affected by Mg^{2+} binding

The addition of magnesium to PhoB RD prior to crystallization resulted in crystal growth under the same conditions as for the apo form, but in a new space group, $C222_1$. In this new crystal form (PhoB RD bound to Mg^{2+}; referred to in the following as $PhoB^{RM}$) there are three molecules per asymmetric unit, Mg-A, Mg-B and Mg-C (Fig. 1*a*), all of which contain an Mg^{2+} ion in the active site. Each monomer shows the PhoB RD doubly wound $(\alpha/\beta)_5$-fold consisting of a five-stranded twisted parallel β-sheet of topology $\beta2\beta1\beta3\beta4\beta5$, flanked by two helices ($\alpha1$ and $\alpha5$) on the concave side and three ($\alpha2$, $\alpha3$ and $\alpha4$) on the convex side (Fig. 1*b*). Strands and helices alternate along the primary sequence and $\beta1$, $\beta3$ and $\beta4$ form the protein core. The active site is located at the C-terminal edge of the central β-sheet, in a wide cavity formed by loops connecting strands $\beta1$, $\beta3$ and $\beta5$ with their subsequent helices $\alpha1$, $\alpha3$ and $\alpha5$ (L$\beta1\alpha1$, L$\beta3\alpha3$ and L$\beta5\alpha5$, respectively). The last residue of $\beta3$, Asp53, is the invariable phosphorylation site (Makino *et al.*, 1989; Hiratsu *et al.*, 1995). Glu9, Asp10 and Glu11 forms a conserved acidic triad in the active site (see §3.4) that together with Asp53 confer a net negative charge on the cavity, which is extended by residues of the nearby segments L$\beta2\alpha2$ (Glu33 and Asp34), the N-terminus of helix $\alpha2$ (Asp36) and helix $\alpha4$ (Glu87, Glu88, Glu89 and Asp90). The fold in PhoB RD is conserved among other RDs belonging to the OmpR/PhoB subfamily, such as PhoP (PDB code 1mvo; Birck *et al.*, 2003), DrrB (PDB code 1p2f; Robinson *et al.*, 2003), DrrD (PDB code 1kgs; Buckler *et al.*, 2002), $Rr02^{R}$ (also termed MycA; PDB codes 1nxt and 1nxo; Bent *et al.*, 2004), $ArcA^{R}$ (PDB codes 1xhe and 1xhf; Toro-Román, Mack *et al.*, 2005), $KdpE^{R}$ (PDB codes 1zh2 and 1zh4; Toro-Román, Wu *et al.*, 2005) and $TorR^{R}$, (PDB code 1zgz; Toro-Román, Wu *et al.*, 2005), but also in more distant relatives RDs such as CheY (Stock *et al.*, 1989), CheB (Djordjevic *et al.*, 1998; Djordjevic & Stock, 1998), NtrC (Volkman *et al.*, 1995), Spo0F (Madhusudan *et al.*, 1996), NarL (Baikalov *et al.*, 1996), FixJ (Gouet, Fabry *et al.*, 1999), DctD (Park *et al.*, 2002), Spo0A (Lewis *et al.*, 1999), ETR1 (Muller-Dieckmann *et al.*, 1999) and DivK (Guillet *et al.*, 2002).

Of the three molecules in $PhoB^{RM}$, only Mg-B displays continuous electron density, whereas Mg-A and Mg-C show discontinuities at helix $\alpha4$, making it difficult to trace the polypeptide chain. The root-mean-square deviation (r.m.s.d.) between Mg-A and Mg-B is 0.35 Å (for 110 common C^{α} atoms deviating less than 3 Å), with maximal differences located at the ends of the loops flanking Mg-B helix $\alpha4$ (Gly86 C^{α}, 2.6 Å; Leu95 C^{α}, 1.6 Å). Mg-B and Mg-C (r.m.s.d. of 0.27 Å for 108 C^{α} atoms) also deviate maximally in the same area (Ala84, 0.5 Å; Thr97, 1.8 Å). Mg-A and Mg-B establish the same dimer as described for monomers WT-A and WT-B in the apoprotein (PDB code 1b00; Solà *et al.*, 1999; the r.m.s.d. between dimers is 0.61 Å), as they are related by an almost perfect non-crystallographic twofold axis to which both β-sheet C-termini converge (Fig. 1*c*), joining the two active sites in a large electronegative depression (Fig. 1*e*). Underneath and perpendicularly, the dimerization surfaces are self-complementary and glued by a hydrophobic core. The contact regions include helix $\alpha1$, the loop L$\beta5\alpha5$ and the N-terminus of helix $\alpha5$ from both monomers (the $\alpha1$-L$\beta5\alpha5$ interface). Between Mg-A and Mg-B the $\alpha1$-L$\beta5\alpha5$ interface spreads over 578 $Å^2$ and the protomers are almost perfectly superimposable onto the WT-A/B dimer (r.m.s.d. of 0.61 Å; Fig. 1*c*). A stereochemical analysis of the interface area reveals that its size is just below the mean found for transient homodimer interaction surfaces (740 $\pm$ 139 $Å^2$; Nooren & Thornton, 2003). However, the rather small interaction area is compensated by a higher number of non-polar atoms (74% in both Mg-A and Mg-B) than the average for this kind of surface (65 $\pm$ 7%; Nooren & Thornton, 2003). The residues involved are Ile14, Met17, Phe20, Val21, Gln24 (all in $\alpha1$), Pro106 (L$\alpha1\beta1$) and Pro109 (L$\beta5\alpha5$) from either protomer, plus Pro13 ($\alpha1$) from Mg-A (Fig. 1*d*). The edge of the interface bordering the

active site includes conserved positions within the PhoB/OmpR family (Ile14, 59% and Pro106, 100%), whereas the core of the interaction is non-conserved (0% similarity; Fig. 1*d*). The same interaction appears again in the crystal *via* a crystallographic dyad between Mg-C and Mg-C′, a symmetry-related molecule (Fig. 1*c*; deviation with respect to the WT dimer is 0.55 Å). In this case, the surface area is 544 Å^2 and the symmetric participants are the same, except that Gln24 is present and Phe107 is missing. In the WT-A/B dimer (560 Å^2 interacting surface), the same residues are observed, except that Phe107 and Pro13 are present and Gln24 is missing in both protomers. Note that the variable residues are limiting the interface, so their presence depends on small rearrangements between protomers in different dimers. Although this α1-Lβ5α5 interface has not been reported in any other dimeric RD, it has been found systematically in almost all crystals of PhoB, namely the two apo-PhoB RD structures reported in different crystallization conditions (Solà *et al.*, 1999; Bachhawat *et al.*, 2005), PhoB^{RM} (the present study) and two double-mutant structures (with similar stereochemical values of planarity and polarity, unpublished results). This interface is formed by both crystallographic and non-crystallographic dyads and it is present in crystals of different space groups, such as $P2_12_12_1$ (Solà *et al.*, 1999), $P6_122$ (Bachhawat *et al.*, 2005), $P2_1$ and $C222_1$ (unpublished results and this work). This specific dimer has been associated with the inactive state of PhoB by solution studies on apo-PhoB RD using analytical centrifugation, dynamic light-scattering and NMR measurements at different concentrations in absence of Mg^{2+} (Bachhawat *et al.*, 2005). These studies indicated that these dimers are weak and are in equilibrium with monomers. All the present data suggest that PhoB has a strong tendency to preform the α1-Lβ5α5 interaction during crystallization, probably induced by the high protein concentration, which displaces the equilibrium to the non-active dimeric form.

3.2. The PhoB RD–Mg^{2+} complex shows an incipient active interface (the 'Pre-α4-β5-α5 interface')

In the PhoB^{RM} crystal form, Mg-B and Mg-C make a second type of interaction which involves helix α4 of Mg-B and β5 and helix α5 of both molecules (referred to in the following as the Pre-α4-β5-α5 interface; see Fig. 1*a*). The prefix Pre- is included as the interacting surface can be envisaged as an intermediate state towards the final activated one (see below). The Pre-α4-β5-α5 interaction surface buries only 382 Å^2 per monomer, contains more than 40% of polar residues and is not perfectly symmetric. Only three residues from each of Mg-B and Mg-C are directly involved in the interaction: Asp101 (β5), Arg115 (α5) and Arg122 (α5) (Figs. 2*a* and 2*b*). Mg-B makes additional interactions with the carbonyl O atom of Asp100 (β5) and the side chains of Asp90 (α4), Tyr102 (β5) and Arg121 (α5) and Mg-C contributes the carbonyl groups of Ala99 (Lβ4α5) and Tyr102 (β5) and the side chains of Lys110 and Glu111 (α5). The high flexibility of helix α4 in Mg-C further contributes to the asymmetry of this interface, as the residues therein do not take part in the contacts (Fig. 2*b*). However, although all parameters point to a nonspecific crystal interaction, three features are of special interest. Firstly, the same segments belong to the dimerization interface of the fully activated RD of PhoB (PhoB^{RAc}), the active site of which (Fig. 2*f*) contains an Mg^{2+} atom and a phosphoryl analogue, beryllium fluoride (BeF_3^-; Yan *et al.*, 1999; Bachhawat *et al.*, 2005). The PhoB^{RAc} contacting area is symmetric and because it involves all three segments, helix α4, β5 and helix α5 from both molecules (α4-β5-α5 interface), is much larger (890 Å^2 per monomer). Secondly, the Mg-B and Mg-C interface contains residues that are highly conserved within the OmpR/PhoB family (Fig. 2*a*). Thirdly, RD structures of close relatives, such as the OmpR/PhoB family members *E. coli* ArcA^R (Toro-Román, Mack *et al.*, 2005; free or in complex with Mg^{2+} and BeF_3^-) and the non-activated pneumococcal MicA^R (Bent *et al.*, 2004), have a very similar dimerization interface (850 and 995 Å^2, respectively) to PhoB^{RAc}. Further homologues with a similar interface include KdpE^R and TorR^R (Toro-Roman, Wu *et al.*, 2005). It has been suggested that a similar oligomerization arrangement upon phosphorylation may occur for all OmpR/PhoB family members (Bachhawat *et al.*, 2005). A special feature of both the Pre-α4-β5-α5 and full α4-β5-α5 interactions is that the two active sites within the dimer are separated by the interface segments and placed on opposite faces (Fig. 2*c*), unlike the Mg-A/B α1-Lβ5α5 oligomer active sites, which lie within the same cavity (Fig. 1*e*).

The PhoB^{RAc} structure shows three molecules, Be-A, Be-B and Be-C, in the asymmetric unit, which have BeF_3^- and a Mg^{2+} ion at the active site (Bachhawat *et al.*, 2005). Interestingly, Be-C makes contact with a symmetrically related molecule A′ (Be-A′) through its α1-Lβ5α5 surface; however, Be-A′ is rotated 90° with respect to Mg-A. Be-B and Be-C interact through the α4-β5-α5 interface, but several specific features differentiate it from the Pre-α4-β5-α5 interaction seen in activated PhoB^{RM}. The two PhoB^{RM} B/C and PhoB^{RAc} C/B dimers superpose with an r.m.s.d. of 3.31 Å. The deviations are generated by a relative rotation of 20° of the respective subunits and because the corresponding interacting segments (α4, β5 and α5) are differently organized. The protomers Mg-B and Be-C (r.m.s.d. of 1.73 Å for 120 C$^\alpha$ atoms) are equivalent in the dimer but deviate essentially at the Lβ4α4 and helix α4 segments (if these segments are excluded, the r.m.s.d. reduces to 0.63 Å for 105 C$^\alpha$ atoms). If we consider the PhoB^{RM} structure as the Mg^{2+}-bound precursor of PhoB^{RAc}, we can see that upon BeF_3^- binding the peptide bond between Ala84 and Arg85 (at Lβ4α4) flips and the Arg85 side chain moves from the molecular surface to above the cleft between helices α3 and α4 and lies above the aromatic ring of Trp54. Moreover, helix α4 not only elongates (from Val92 in Mg-B to Thr97 in Be-C) but also rotates; therefore, the residues exposed in PhoB^{RM} become oriented towards the protein core in PhoB^{RAc} and *vice versa*. One example is Glu88, which interacts with the active site in Mg-B or WT-A (Solà *et al.*, 1999; Fig. 2*d*) but is driven to the solvent after the helical shift (Fig. 2*f*), or Asp90, which upon phosphorylation loses a salt bridge with Mg-C Lys110 (helix α5) at

the PhoB^{RM} dimer interface (Fig. 2*b*). Val92 is the only hydrophobic residue that anchors helix α4 to the protein core either in apo-PhoB RD or in PhoB^{RM}. Val92 has been transported to the dimer interface by the helical rotation in PhoB^{RAc}, where it forms part of the interaction patch on the α4-β5-α5 surface, which is symmetric, well defined and includes hydrophobic and hydrophilic contacts in PhoB^{RAc} (Bachhawat *et al.*, 2005). The RD structures of the PhoB orthologues *Streptococcus pneumoniae* MicA^{R} (Bent *et al.*, 2004) and *E. coli* ArcA^{RAc} (Toro-Román, Mack *et al.*, 2005), KdpE^{RAc} and TorR^{R} (Toro-Román, Wu *et al.*, 2005), display the α4-β5-α5 PhoB^{RAc} interactions, with helix α4 also rotated. Positions in this α4-β5-α5 conserved patch of interactions include invariable residues such as Asp101 (100% conserved; Fig. 2*a*) and Arg115 (100% conserved; Fig. 2*a*). In PhoB^{RAc} these polar residues establish perfectly symmetric intermolecular salt bridges at the core of the interface. However, although these residues also meet at the Pre-α4-β5-α5 interface in PhoB^{RM}, their interactions are not symmetric. Arg115 residues from either molecule make a stacking interaction through their guanidinium group (Fig. 2*b*). Such an interaction is not possible in PhoB^{RAc} owing to the relative rotation of the protomers with respect to PhoB^{RM}. Likewise, Arg122 from both Mg-B and Mg-C also interact and are further bound to the carbonyl O atom of Mg-B Asp101 (Fig. 2*b*). In PhoB^{RAc}, each of these arginines make a salt bridge with the Asp100 carboxylate from the opposite protomer. The PhoB^{RAc} structure shows, like all the other activated RDs (*e.g.* Bent *et al.*, 2004; Birck *et al.*, 1999; Hastings *et al.*, 2003; Lee, Cho, Pelton, Yan, Berry *et al.*, 2001; Lee, Cho, Pelton, Yan, Henderson *et al.*, 2001; Toro-Román, Mack *et al.*, 2005; Toro-Román, Wu *et al.*, 2005), Tyr102 (at β5) in a *trans* (*t*) rotamer. The side chain is pointing towards Lβ4α4, a movement allowed by the *t* rotamer of Thr83 (end of β4), which binds to one of the phosphoryl O atoms at the active site. These related orientations have been termed the YT coupling (Lee, Cho, Pelton, Yan, Berry *et al.*, 2001) and imply respective *t*/*t* conformations of both Thr83/Tyr102. In PhoB^{RM}, Tyr102 is located at the α4-β5-α5 interface (Figs. 2*a* and 2*b*). None of the PhoB^{RM}

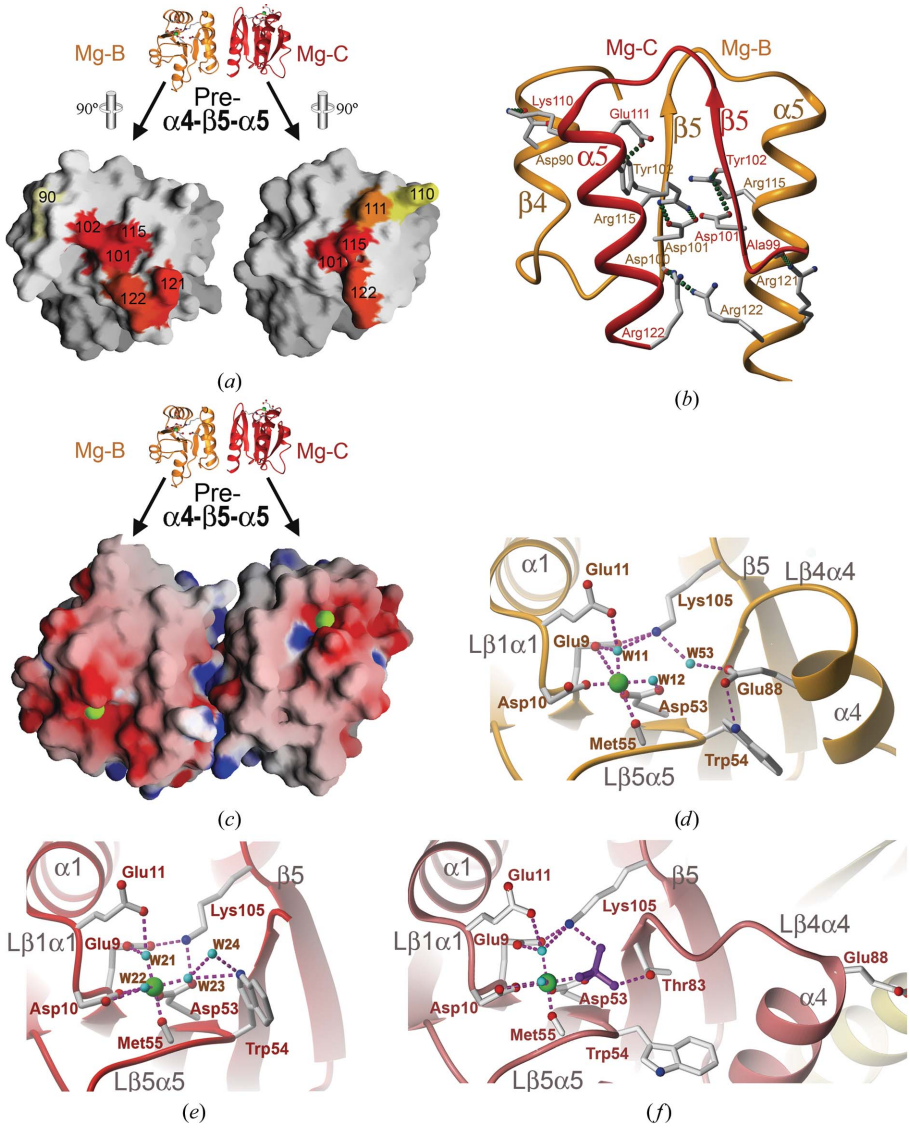


Figure 2
The Pre-α4-β5-α5 interface. (*a*) The α4-β5-α5 interface is made up by Mg-B (left, orange) and Mg-C (right, red). Each constituent monomer (superimposed on its Connolly solid surface) has been rotated vertically 90° from the position occupied in the small Richardson diagram to allow insight into the interacting surface. The residues of each molecule participating in direct contacts are shown and labelled. The colour coding, ranging from 0% (light yellow) to 100% (intense red), reflects the conservation of each position after sequence alignment. (*b*) Detail of the α4-β5-α5 interface showing the side chains of the residues engaged in direct contacts from Mg-B (orange Richardson diagram) and Mg-C (red). Hydrogen bonds and salt bridges are characterized by dashed lines. (*c*) Representation of the Connolly surfaces of Mg-B and Mg-C protomers in PhoB^{RM}. The green spheres represent the two Mg^{2+} ions in the respective active sites, which are inside independent cavities orientated towards different faces. The two Mg-B and Mg-C protomers contact by their α4-β5-α5 surfaces, although the interaction is not fulfilled (therefore, it is a Pre-α4-β5-α5 interaction) as in activated PhoB (Fig. 2*c*; Bachhawat *et al.*, 2005). (*d*) Close-up view of the active site of Mg-B. Only interactions involving side-chain atoms (except for Met55 O), the magnesium cation (green sphere) and solvent molecules are shown. Secondary-structure elements are indicated. Note the presence of Glu88 inside the cavity and the location of both the Trp54 side chain and the helix α4 segment. (*e*) Close-up view of Mg-C active site with the same representation criteria as for (*d*). Trp54 directly participates in the active-site network of interactions, while helix α4 is not organized. (*f*) Close-up view of the Be-B active site activated with $BeF3^{-}$, represented in violet (Bachhawat *et al.*, 2005). Helix α4 is shifted with respect to the $BeF3^{-}$-free molecule (*d*) and therefore Glu88 is orientated towards the surface of the domain.

protomers show the YT coupling, as the Thr83/Tyr102 conformations are *gauche*$^-$/*gauche*$^-$ (g^-/g^-) (in Mg-A), g^+/g^+ (Mg-B) and g^-/g^- (Mg-C). In Mg-C, helix α4 is disordered from Arg85 to Glu96, presumably abrogating the intramolecular interaction of Tyr102 with Arg93, thus inducing the g^- conformation of Tyr102.

In summary, comparison of the PhoB^{RM} and PhoB^{RAc} structures shows that PhoB RD crystallized in complex with Mg^{2+} results in an increase of the flexibility of the α4-β5-α5 surface that promotes a weak protein–protein contact in the crystal which resembles the final active interaction.

3.3. Trp54 side-chain orientation depends on helix α4 stability

Trp54 is another interesting residue that fluorescence studies have identified as an indicator of PhoB activation (Zundel *et al.*, 1998). It is placed at the beginning of Lβ3α3, between two important positions: the phosphorylable Asp53 and the metal-coordinating Met55 (Fig. 2*d*; see §4). In the WT-A and WT-B monomers, Trp54 is in a *trans* conformation, the side chain fitting snugly in a hydrophobic cavity between helices α3 and α4. Interestingly, in Mg-B the whole last turn of helix α4 unwinds from Arg93 on; Leu95 C$^\alpha$, for example, is more than 5 Å away from its position in WT and is exposed to the solvent. Consequently, the cavity that surrounds Trp54 in apo-PhoB disappears in Mg-B. This residue rotates 30° around its χ^1 angle, bringing the indole ring close to the active site, where its N$^{\varepsilon1}$ hydrogen binds to Glu88 O$^{\varepsilon2}$ (Fig. 2*d*). The same correlation between rearrangements of helix α4 and Trp54 rotamers is found in both Mg-A and Mg-C, where the instability of the entire helix α4 cancels a whole side of the Trp54 hydrophobic pillow. In Mg-A, Trp54 is poorly defined but residual density suggests that the side chain oscillates from *t* to g^- conformations, alternating the position between the helices and the active-site edge. In Mg-C, the side chain of Trp54 rotates further than in Mg-A and penetrates into the active-site rim, where it is surrounded by Met55 (β3) and the main chain of Thr83 and Ala84 (β4 end) (Fig. 2*e*). Indeed, its N$^{\varepsilon1}$ atom interacts with W23 from the magnesium coordination sphere (see §4 and Figs. 1*f* and 2*e*). The area that Trp54 side chain occupies in Mg-C is the same as that filled by Glu88 in both WT protomers and Mg-B (Fig. 2*d*), which is empty in Mg-C owing to the disorder of helix α4 (Fig. 2*e*).

3.4. Magnesium binding to the active site of PhoB RD reveals asymmetry inside the catalytic cavity

The activated structure shows that the magnesium ion and the phosphate moiety (or its emulator BeF_3^-) locate contiguously within the active site (see Fig. 2*f*; Birck *et al.*, 1999; Hastings *et al.*, 2003; Toro-Román, Mack *et al.*, 2005; Bachhawat *et al.*, 2005). We define two adjacent areas in the cavity: site M (for Mg^{2+}) and site P (for phosphate). Site M is surrounded by Lβ1α1, which includes residues involved in direct or indirect coordination to the Mg^{2+} atom, such as the highly conserved triad Glu9, Asp10 and Glu11. At the other end of the cavity, β4 delimits site P, where Thr83 O$^\gamma$ and Ala84 N are involved in phosphoryl oxygen (or BeF_3^- fluorine) binding (Fig. 2*f*). The P and M sites are bridged by the side chain of Asp53, which becomes phosphorylated on its O$^{\delta1}$ atom during activation and simultaneously binds the Mg^{2+} cation through its O$^{\delta2}$ atom. The following backbone of Lβ3α3 is also partitioned, with the amide N atoms of positions 54 and 55 binding the phosphate moiety and the carbonyl O atom of residue 55 coordinating the cation. On the other side, Lys105 N$^\zeta$ (a nearly invariant residue) from Lβ5α5 binds a phosphate O atom but contacts site M *via* the carboxylate of Glu9. Comparison of apo-RD structures of PhoB, ArcA, CheY, FixJ and DctD with their activated variants clearly shows that upon binding, the rigid *cis*-Pro-containing main chain adjacent to Lys105 transmits the movement to the neighbouring C-terminal end of β4 and Lβ4α4 (Santoro *et al.*, 1995; Bellsolell *et al.*, 1996).

In order to compare the Mg^{2+}-bound active sites with each other and with the apo-PhoB RD molecules, only loops Lβ1α1 and Lβ5α5 were superimposed, as it is only around the M site that structures are rigid and comparable. Accordingly, the superposition of Mg-A, Mg-B and Mg-C molecules shows that the magnesium cations cluster at identical positions in all three active sites (differences are less than 0.2 Å). For the metal assignment at the density found at site M, it was taken into account that Mg^{2+}, Na^+ and solvent molecules are isoelectronic and therefore the distinction between them cannot be based on electron density or comparison of the temperature factors of surrounding atoms. Both sodium and magnesium typically display an octahedral coordination by six oxygen ligands, with a theoretical cation–oxygen distance of 2.46 Å for Na^+ and 2.07 Å for Mg^{2+} (Harding, 1999, 2002). Mg-C shows a complete octahedral coordination sphere for the cation (Fig. 1*f*), including Asp10 O$^{\delta1}$ (2.2 Å from Mg^{2+}), Asp53 O$^{\delta2}$ (1.9 Å) and Met55 O (2.2 Å). Solvent molecules W21 (2.0 Å), W22 (2.1 Å) and W23 (2.2 Å) complete the coordination. The same atoms coordinate the metal ion in Mg-A and Mg-B except for the W22 equivalents, which are missing in these two molecules (Fig. 2*d*). The distances to the cation in Mg-A (values for Mg-B are given in parentheses) are Asp10 O$^{\delta1}$ 2.0 Å (2.0 Å), Asp53 O$^{\delta2}$ 2.1 Å (1.8 Å), Met55 O 2.5 Å (2.8 Å), W1 2.0 Å (W11 2.1 Å) and W2 2.1 Å (W12 2.1 Å). Agreement between the three protomers is also observed for the side chains of Glu9, Asp10, Glu11 and Asp53 (Figs. 2*d* and 2*e*). While Asp10 coordinates the magnesium ion, Glu9 and Gl11 interact with the metal through one of the apical coordinating water molecules (W2, W11 and W21 for Mg-A, Mg-B and Mg-C, respectively; Figs. 2*d* and 2*e*). In parallel, Glu9 establishes a highly conserved salt bridge with Lys105 N$^\zeta$ (Solà *et al.*, 1999; Welch *et al.*, 1994; Park *et al.*, 2002; Toro-Román, Mack *et al.*, 2005; Bachhawat *et al.*, 2005; Toro-Roman, Wu *et al.*, 2005), which shows good overlap (maximal deviation 0.4 Å) upon superposition of the Mg^{2+}-bound structures shown here. This interaction is postulated (Solà *et al.*, 1999) to fix the lysine N$^\zeta$ in a position suitable for a further salt bridge with Asp53 O$^{\delta2}$, which is expected to be disrupted during the activation events. This leaves the Asp53 carboxylate oxygen free to act as the attacking nucleophile in the

phosphotranspher reaction, as shown in PhoB^{RAc} (Bachhawat *et al.*, 2005). Furthermore, the Glu9–Lys105 salt bridge is complemented by a solvent bridge between Lys105 and Glu11 found in Mg-A, Mg-B and WT-B. Another systematic feature is the mentioned apical position of the ion coordination sphere, occupied in the three PhoB^{RM} protomers and also appearing in the metal-free WT-A molecule. In Mg-B, the Mg^{2+}-coordinating solvent molecules W12 and W53 extend the network of interactions to the rest of the active-site cavity (Fig. 2*d*). Overall, site M conserves the positions of the metal-coordinating protein atoms and the apical solvent molecules and some interactions appear very often, such as those established by Lys105 with Glu9, Glu11 and Asp53.

It is interesting that neither water molecules nor side chains are as prevalent in site P as in site M. In Mg-C, the Trp54 side chain penetrates the active site (see §3.3) and interacts with the metal-coordinating solvent molecule W21 and with W24 (Figs. 1*f* and 2*e*). In contrast, in Mg-B this side chain is located outside the cavity and the P site is filled by the carboxyl O atoms of Glu88, which can thus interact (Glu88 $O^{\varepsilon2}$—Trp54 $N^{\varepsilon1}$, see Fig. 2*d*). Glu88 $O^{\varepsilon1}$ also interacts with Ala84 N (at $\beta4$) and with another solvent molecule, W53, which links the glutamate to site M *via* Lys105 and the phosphorylable Asp53 (Fig. 2*d*). Glu88 is also present in the catalytic cavity in both WT-A and WT-B monomers, interacting with Ala84 N. Furthermore, in WT-A it forms a water bridge to Asp53 *via* the solvent molecule W53. Interestingly, this water molecule occupies the position of Trp54 $N^{\varepsilon1}$ in Mg-C. In Mg-A, Glu88 is absent from the P site owing to the disorder of helix $\alpha4$, but residual electron density suggests that the Trp54 indole ring partially penetrates the cavity, as mentioned in §3.3. Summarizing, PhoB^{RM} shows that the interactions appear to be regular at the M site, but there is a mutually exclusive presence of Glu88 and Trp54 at the P site. Both residues can make interactions inside the cavity. As explained above, the presence of Glu88 in the active site fixes helix $\alpha4$, whilst the presence of Trp54 correlates with the destabilization of this helical segment. Taken together, our results show that upon Mg^{2+} binding, Glu88 may be substituted by Trp54 at the active site, resulting in the removal of an anchor of helix $\alpha4$ with the protein core, thus promoting an unstructured state that may predispose this segment for an eventual shift at the final step of activation.

3.5. Comparison of Mg^{2+}–PhoB^{RM} and BeF_3^-– PhoB^{RAc} active sites

The superposition of the active sites of Mg-B and the activated PhoB^{RAc} molecule Be-C shows that the Mg^{2+} ion and the atoms from the coordinating magnesium sphere fit very well (data not shown). Interestingly, the Mg-B metal-coordinating W12 coincides (at 0.4 Å distance) with the position of one of the BeF_3^- F atoms, F1, which also coordinates the Mg^{2+} (Figs. 2*d* and 2*f*). In PhoB^{RAc} the F2 and F3 atoms of BeF_3^- are coordinated by the protein atoms Thr83 O^{γ} and Ala84 N, respectively (Bachhawat *et al.*, 2005). Be, F2 and F3 form a plane and the central solvent molecule W53 of Mg-B is close to the F2–F3 edge of this plane. However, W53 does not bind to Ala84 N but rather to Glu88 $O^{\varepsilon1}$, which in turn is engaged in a hydrogen bond with Ala84 N (Fig. 2*d*). Therefore, the Mg-B structure shows that binding of magnesium in the active site leads to an interaction between solvent molecules (W12 and W53) that emulates the molecular linkage between the fluorines in BeF_3^- (Figs. 2*d* and 2*f*). In Mg-A and Mg-C the presence of these interactions is prevented by the entrance of the Trp54 side chain, which is concomitant with the absence of Glu88 carboxylate owing to the disorder of helix $\alpha4$ (Fig. 2*e*). In PhoB^{RAc}, Thr83 O^{γ} of Be-C, which coordinates the F2 atom (Fig. 2*f*), is closer (4.7 Å difference) to the active site than in Mg-B, owing to movement of $\beta4$ and the rotation of the Thr83 side chain from g^- to t conformation in the former molecule. The amide Ala84 N of both Mg-B and Be-C are 1.8 Å apart owing to the movement of $\beta4$ in PhoB^{RAc} towards the active site as it coordinates F3. Comparison of the PhoB^{RM} and PhoB^{RAc} structures suggests that once the Mg^{2+} ion is bound to the PhoB active site, the subsequently bound phosphate group (or BeF_3^- molecule) substitutes both the metal-bound W12 and the W53 solvent molecules and drags the main chain of $\beta4$ towards the active site, including both Thr83 and Ala84. Both events, the substitution of the solvent molecules and the $\beta4$ movement, impede the interactions that Glu88 or, in its absence, Trp54, establishes with the Mg^{2+}-containing active site before phosphorylation (Figs. 2*d*, 2*e* and 2*f*).

4. Discussion

4.1. Binding of the Mg^{2+} atom to PhoB RD promotes a transitory flexible conformation of helix $\alpha4$

Active-site analysis of unbound PhoB RD and its complex with Mg^{2+} shows that there are two differentiated areas inside the cavity, the M site and the P site. The M site, framed by L$\beta1\alpha1$ and L$\beta3\alpha3$, harbours in all structures a network of interactions that persists upon addition of magnesium. In contrast, the P site shows a heterogeneous organization. In Mg-B, the solvent molecule W53 connects the network from the M site to the P site (Figs. 1*f* and 2*d*). The interactions of solvent molecules W12 and W53 in the Mg-B active site are similar to the contacts established by the activating moiety (BeF_3^-) in PhoB^{RAc} (Figs. 2*d* and 2*f*, respectively). Glu88 in site P participates in this network (Fig. 2*d*). In contrast, in the P sites of Mg-A and Mg-C neither solvent molecule W53 nor Glu88 is found (Fig. 2*e*). Instead, the Trp54 indole moiety is present and in Mg-C it participates in the network of active-site interactions (Fig. 2*e*). In apo-PhoB^{R}, the Trp54 side chain is found between helices $\alpha3$ and $\alpha4$, which show different relative positions in WT-A and WT-B. However, despite these differences, the Trp54 rotamer is not altered. Rather, the versatile orientation of the indole moiety follows the conformational changes of the highly flexible helix $\alpha4$. For example, in Mg-B the disorder of helix $\alpha4$ at the C-terminus correlates with the position of the Trp54 side chain, which changes its van der Waals interaction partners from the lower to the upper part of the helical segment and makes an interaction with

Glu88 (Fig. 2*d*). In Mg-A, and especially in Mg-C, helix α4 is entirely disordered and Trp54 is immersed in the active-site rim, filling the space liberated by Glu88 (Fig. 2*e*). All these facts suggest a working model hypothesis. Upon Mg^{2+}-complex formation, solvent molecules generate a new set of interactions in the active site occupying the position of the phosphate moiety in the activated form (Fig. 2*d*). Glu88 is bound to this network (Fig. 2*d*). However, these interactions are interrupted by the Trp54 side chain, which displaces Glu88 from the active site and concomitantly destabilizes helix α4 (Fig. 2*e*). Trp54 and Glu88 may alternate at the active site, with related intermittent disordered/ordered conformations of helix α4. Therefore, the presence of Trp54 in the active site contributes to an increase of the degrees of freedom of helix α4, which allows this segment to rotate. Once phosphorylation at Asp53 $O^{\delta1}$ occurs, Trp54 is displaced from the active site and Thr83 is attracted to the cavity, which allows Tyr102 to adopt the active *t* conformation. $PhoB^{RAc}$ shows that Trp54 adopts a rotamer that fits the new conformation (elongated and shifted) of helix α4. Interestingly, CheY cocrystallized in the presence of Mg^{2+} (lacking the activating phosphate) shows a distortion of Lβ4α4 and an unwinding of helix α4 (Bellsolell *et al.*, 1994). Other cases where non-activated and activated RD structures can be compared include FixJ (Birck *et al.*, 1999), DctD (Park *et al.*, 2002) and NtrC (Hastings *et al.*, 2003). Those show that residues at the position equivalent to Trp54 (Leu55 in FixJ, Ile56 in DctD and Ile55 in NtrC) undergo side-chain or even main-chain rotation to accommodate the YT coupling (involving equivalents of Thr83 and Tyr102), which also correlates with movements in helix α4. In $PhoB^{RAc}$, the side chain of Trp54 is within van der Waals distance of Thr83 and Tyr102 coupled residues, which are in *t* conformations. According to all these data, inclusion of magnesium at the active site increases the interactions inside the cavity, but also induces an increase of flexibility of helix α4 that facilitate a conformational change of this segment, which is stabilized by phosphorylation. The movement of Trp54 as it enters the active site is a key feature in the process of PhoB activation. The full-length DrrB structure shows that destabilization of helix α4 could also lead to a release of the ED (Bachhawat *et al.*, 2005).

4.2. The inactive α1-Lβ5α5 interface is systematically present in all studied PhoB RD structures

It was suggested that ablation of the C-terminal ED would eliminate the inactivating steric hindrance between domains and establish the functional RD active dimer, even in the absence of activating events (Fiedler & Weiss, 1995). Except for the fully activated $PhoB^{RAc}$ (Bachhawat *et al.*, 2005), all the ED-depleted PhoB RD dimers studied are joined by the α1-Lβ5α5 interface, despite the crystals belonging to distinct space groups and the respective dyads of the dimers being either crystallographic or non-crystallographic. Indeed, preliminary structural studies on two constitutive double-point mutant RDs of PhoB RR further confirm this finding (unpublished results). Therefore, this interface appears in isolated PhoB RDs and in the absence of activating events and is reproducible in different crystal contexts. Biophysical studies indicated that it corresponds to a weak interaction (Bachhawat *et al.*, 2005) promoted in the crystals perhaps because of the high protein concentration. The sequence analysis shows low conservation of the residues involved, suggesting that this interaction is specific to PhoB. However, this interface is distinct from that formed after PhoB RD activation by the phosphate analogue BeF_3^-, as the $PhoB^{RAc}$ structure shows (Bachhawat *et al.*, 2005). A recent model (Bachhawat *et al.*, 2005) proposed that inactive PhoB is in an equilibrium between the α1-Lβ5α5-dimer and the monomeric forms and that phosphorylation of the monomer PhoB would lead to the active α4-β5-α5 dimer. This model is consistent with Fiedler & Weiss (1995) as well as with the dimers observed before and after phosphorylation (McCleary, 1996), which would correspond to two different oligomer forms. Our results show that the α1-Lβ5α5 interface survives magnesium binding in this two-step activation. This suggests that the presence of magnesium does not affect the equilibrium between inactive α1-Lβ5α5-dimers and monomers, although helix α4 undergoes a transition between ordered and destabilized states triggered by Trp54. This might be the usual state of the PhoB receiver domain considering the intracellular levels of soluble Mg^{2+} in bacteria (4 m*M*; Lusk *et al.*, 1968). Our results also show the active conformation is initiated on the α4-β5-α5 surface (Pre-α4-β5-α5) and establishes a quasi-symmetric contact inside the crystal. Furthermore, inactive α1-Lβ5α5 and Pre-α4-β5-α5 interactions are compatible with each other. However, an *in vivo* α1-Lβ5α dimer–dimer interaction through Pre-α4-β5-α5, although conceivable, is unlikely because the corresponding stereochemical parameters suggest that the Pre-α4-β5-α5 contact is very weak. Analysis of the activated $PhoB^{RAc}$ structure shows that the Be-C protomer and a symmetrically related molecule Be-A′ contact each other through their respective α1-Lβ5α5 surfaces, although these are rotated by 90° relative to each other compared with $PhoB^{RM}$. This rotated α1-Lβ5α5 interaction may be required for crystal formation. However, it could also be the result of phosphorylation on the α1-Lβ5α5 interaction. Moreover, in another related RD, Spo0F, this α1-Lβ5α5 surface and the loops from the β-sheet C-terminus are involved in the interaction with the cognate kinase phosphotransferase domain (Zapf *et al.*, 2000). If PhoB had the same kind of interaction with PhoR, this would be an additional molecular strategy to prevent formation of the inactive α1-Lβ5α5 dimer during phosphorylation.

We thank Jenny Colom for help during protein purification and Luis Serrano for support during the initial stages of the PhoB project. We also thank the DESY, ESRF and EMBL local contacts for providing assistance and support for measurements. MS is a beneficiary of the 'Ramon y Cajal' Programme from the Spanish Ministry for Science and Technology. Funding provided by research grants BIO2002-03964, BIO2003-132, GEN2003-20642 and BFU2005-06758/BMC

from the same Ministry, as well as the help provided by the Generalitat de Catalunya (Centre de Referència en Biotecnologia and 2005SGR-00280), are further acknowledged.

References

Aizawa, S.-H., Harwood, C. & Kadner, R. J. (2000). *J. Bacteriol.* **182**, 1459–1471.

Altschul, S. F. & Koonin, E. V. (1998). *Trends Biochem. Sci.* **23**, 444–447.

Bachhawat, P., Swapna, G. V., Montelione, G. T. & Stock, A. M. (2005). *Structure Fold. Des.* **13**, 1353–1363.

Baikalov, I., Schröder, I., Kaczor-Grzeskowiak, M., Gunsalus, R. P. & Dickerson, R. E. (1996). *Biochemistry*, **35**, 11053–11061.

Barrett, J. F. & Hoch, J. A. (1998). *Antimicrob. Agents Chemother.* **42**, 1529–1536.

Bellsolell, L., Cronet, P., Majolero, M., Serrano, L. & Coll, M. (1996). *J. Mol. Biol.* **257**, 116–128.

Bellsolell, L., Prieto, J., Serrano, L. & Coll, M. (1994). *J. Mol. Biol.* **238**, 489–495.

Bent, C. J., Isaacs, N. W., Mitchell, T. J. & Riboldi-Tunnicliffe, A. (2004). *J. Bacteriol.* **186**, 2872–2879.

Birck, C., Chen, Y., Hulett, F. M. & Samama, J. P. (2003). *J. Bacteriol.* **185**, 254–261.

Birck, C., Mourey, L., Gouet, P., Fabry, B., Schumacher, J., Rousseau, P., Kahn, D. & Samama, J. P. (1999). *Structure Fold. Des.* **7**, 1505–1515.

Blanco, A. G., Solà, M., Gomis-Rüth, F. X. & Coll, M. (2002). *Structure Fold. Des.* **10**, 701–713.

Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J.-S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T. & Warren, G. L. (1998). *Acta Cryst.* D**54**, 905–921.

Buckler, D. R., Zhou, Y. & Stock, A. M. (2002). *Structure Fold. Des.* **10**, 153–164.

Carranza, C., Inisan, A.-G., Mouthuy-Knoops, E., Cambillau, C. & Roussel, A. (1999). *AFMB Activity Report 1996–1999*, pp. 89–90. Marseille: CNRS-UPR 9039.

Djordjevic, S., Goudreau, P. N., Xu, Q., Stock, A. M. & West, A. H. (1998). *Proc. Natl Acad. Sci. USA*, **95**, 1381–1386.

Djordjevic, S. & Stock, A. M. (1998). *J. Struct. Biol.* **124**, 189–200.

Ellison, D. W. & McCleary, W. R. (2000). *J. Bacteriol.* **182**, 6592–6597.

Evans, P. (1993*a*). *Proceedings of the CCP4 Study Weekend. Data Collection and Processing*, edited by L. Sawyer, N. Isaacs & S. Bailey, pp. 114–122. Warrington: Daresbury Laboratory.

Evans, S. V. (1993*b*). *J. Mol. Graph.* **11**, 134–138.

Falke, J. J. & Hazelbauer, G. L. (2001). *Trends Biochem. Sci.* **26**, 257–265.

Fiedler, U. & Weiss, V. (1995). *EMBO J.* **14**, 3696–3705.

Forsberg, J., Rosenquist, M., Fraysse, L. & Allen, J. F. (2001). *Biochem. Soc. Trans.* **29**, 403–407.

Gouet, P., Courcelle, E., Stuart, D. I. & Metoz, F. (1999). *Bioinformatics*, **15**, 305–308.

Gouet, P., Fabry, B., Guillet, V., Birck, C., Mourey, L., Kahn, D. & Samama, J. P. (1999). *Structure Fold. Des.* **7**, 1517–1526.

Grefen, C. & Harter, K. (2004). *Planta*, **219**, 733–742.

Guillet, V., Ohta, N., Cabantous, S., Newton, A. & Samama, J. P. (2002). *J. Biol. Chem.* **277**, 42003–42010.

Harding, M. M. (1999). *Acta Cryst.* D**55**, 1432–1443.

Harding, M. M. (2002). *Acta Cryst.* D**58**, 872–874.

Hastings, C. A., Lee, S. Y., Cho, H. S., Yan, D., Kustu, S. & Wemmer, D. E. (2003). *Biochemistry*, **42**, 9081–9090.

Hiratsu, K., Nakata, A., Shinagawa, H. & Makino, K. (1995). *Gene*, **161**, 7–10.

Hoch, J. A. (2000). *Curr. Opin. Microbiol.* **3**, 165–170.

Huber, R. (1965). *Acta Cryst.* **19**, 353–356.

Jones, S. & Thornton, J. M. (1996). *Proc. Natl Acad. Sci. USA*, **93**, 13–20.

Kleywegt, G. J. & Jones, T. A. (1994). *Jnt CCP4/ESF–EACBM Newsl. Protein Crystallogr.* **31**, 9–14.

Koradi, R., Billeter, M. & Wüthrich, K. (1996). *J. Mol. Graph.* **14**, 51–55.

Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). *J. Appl. Cryst.* **26**, 283–291.

Lee, S.-Y., Cho, H. S., Pelton, J. G., Yan, D., Berry, E. A. & Wemmer, D. E. (2001). *J. Biol. Chem.* **276**, 16425–16431.

Lee, S.-Y., Cho, H. S., Pelton, J. G., Yan, D., Henderson, R. K., King, D. S., Huang, L., Kustu, S., Berry, E. & Wemmer D. E. (2001). *Nature Struct. Biol.* **8**, 52–56.

Leslie, A. G. W. (1991). *Crystallographic Computing V*, edited by D. Moras, A. D. Podjarny & J. C. Thierry, pp. 27–38. Oxford University Press.

Lewis, R. J., Brannigan, J. A., Muchova, K., Barak, I. & Wilkinson, A. J. (1999). *J. Mol. Biol.* **294**, 9–15.

Lukat, G. S., Stock, A. M. & Stock, J. B. (1990). *Biochemistry*, **29**, 5436–5442.

Lusk, J. E., Williams, R. J. P. & Kennedy, E. P. (1968). *J. Biol. Chem.* **243**, 2618–2624.

McCleary, W. R. (1996). *Mol. Microbiol.* **20**, 1155–1163.

McCleary, W. R. & Stock, J. B. (1994). *J. Biol. Chem.* **269**, 31567–31572.

Madhusudan, Zapf, J., Whiteley, J. M., Hoch, J. A., Xuong, N.-H. & Varughese, K. I. (1996). *Structure Fold. Des.* **4**, 679–690.

Makino, K., Shinagawa, H., Amemura, M., Yamasa, M. & Nakata, A. (1989). *J. Mol. Biol.* **210**, 551–559.

Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.

Muller-Dieckmann, H. J., Grantz, A. A. & Kim, S.-H. (1999). *Structure Fold. Des.* **7**, 1547–1556.

Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997). *Acta Cryst.* D**53**, 240–255.

Navaza, J. (1994). *Acta Cryst.* A**50**, 157–163.

Nicholls, A., Sharp, K. & Honig, B. (1991). *Proteins*, **11**, 281–296.

Nooren, I. M. & Thornton, J. M. (2003). *J. Mol. Biol.* **325**, 991–1018.

Pao, G. M. & Saier, M. H. Jr (1997). *J. Mol. Evol.* **44**, 605–613.

Park, S., Meyer, M., Jones, A. D., Yennawar, H. P., Yennawar, N. H. & Nixon, B. T. (2002). *FASEB J.* **16**, 1964–1966.

Peränen, J., Rikkonen, M., Hyvonen, M. & Kääriäinen, L. (1996). *Anal. Biochem.* **236**, 371–373.

Robinson, V. L., Wu, T. & Stock, A. M. (2003). *J. Bacteriol.* **185**, 4186–4194.

Santoro, J., Bruix, M., Pascual, J., López, E., Serrano, L. & Rico, M. (1995). *J. Mol. Biol.* **247**, 717–725.

Santos, J. L. & Shiozaki, K. (2001). *Sci. STKE*, **2001**, RE1.

Solà, M., Gomis-Rüth, F. X., Guasch, A., Serrano, L. & Coll, M. (1998). *Acta Cryst.* D**54**, 1460–1463.

Solà, M., Gomis-Rüth, F. X., Serrano, L., Gonzalez, A. & Coll, M. (1999). *J. Mol. Biol.* **285**, 675–687.

Stephenson, K. & Hoch, J. (2002). *Curr. Opin. Pharmacol.* **2**, 507–512.

Stock, A. M., Mottonen, J. M., Stock, J. B. & Schutt, C. E. (1989). *Nature (London)*, **337**, 745–749.

Stock, A. M., Robinson, V. L. & Goudreau, P. N. (2000). *Annu. Rev. Biochem.* **69**, 183–215.

Toro-Román, A., Mack, T. R. & Stock, A. M. (2005). *J. Mol. Biol.* **349**, 11–26.

Toro-Román, A., Wu, T. & Stock, A. M. (2005). *Protein Sci.* **14**, 3077–3088.

Volkman, B. F., Nohaile, M. J., Amy, N. K., Kustu, S. & Wemmer, D. E. (1995). *Biochemistry*, **34**, 1413–1424.

Wang, W., Zhang, W., Chen, H., Chiao, J., Zhao, G. & Jiang, W. (2002). *Arch. Microbiol.* **178**, 376–386.

Wanner, B. L. (1993). *J. Cell. Biochem.* **51**, 47–54.

Wanner, B. L. (1996). *Escherichia coli and Salmonella typhimurium: Cellular and Molecular Biology.*, edited by F. C. Neidhardt, R. I. Curtiss, J. L. Ingraham, E. C. Lin, K. B. J. Low, B. Magasanik, W. Reznikoff, M. Riley, M. Schaechter & H. E. Umbarger, pp. 1357–1381. Washington, DC: American Society for Microbiology.
Welch, M., Oosawa, K., Aizawa, S.-I. & Eisenbach, M. (1994). *Biochemistry*, **33**, 10470–10476.
Wick, L. M. & Egli, T. (2004). *Adv. Biochem. Eng. Biotechnol.* **89**, 1–45.
Wolanin, P. M., Thomason, P. A. & Stock, J. B. (2002). *Genome Biol.* **3**, 1–8.
Yan, D., Cho, H. S., Hastings, C. A., Igo, M. M., Lee, S. Y., Pelton, J. G., Stewart, V., Wemmer, D. E. & Kustu, S. (1999). *Proc. Natl Acad. Sci. USA*, **96**, 14789–14794.
Zapf, J., Sen, U., Madhusudan, Hoch, J. A. & Varughese, K. I. (2000). *Structure Fold. Des.* **8**, 851–862.
Zundel, C. J., Capener, D. C. & MacCleary, W. R. (1998). *FEBS Lett.* **441**, 242–246.